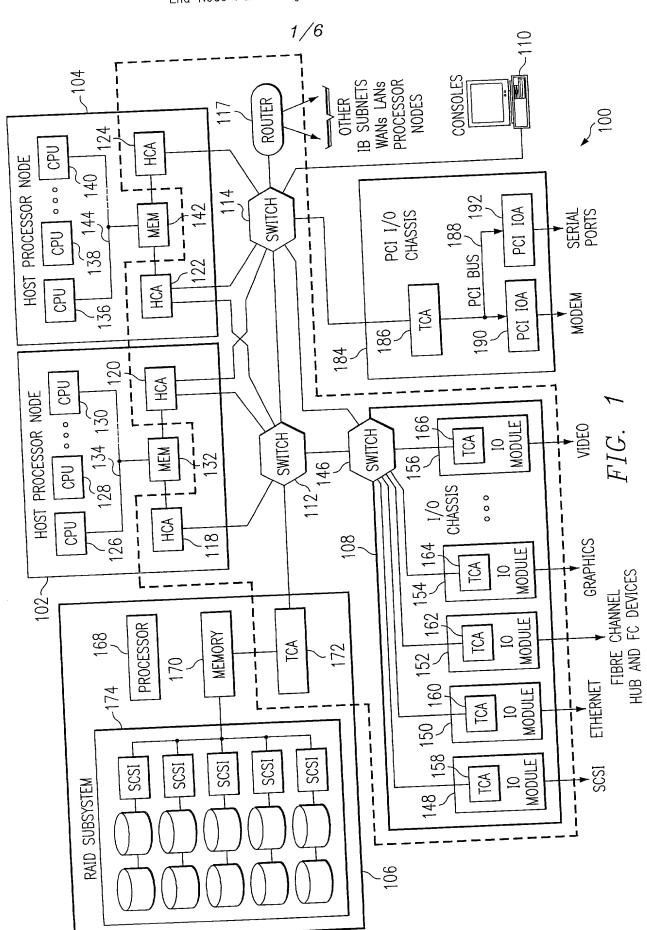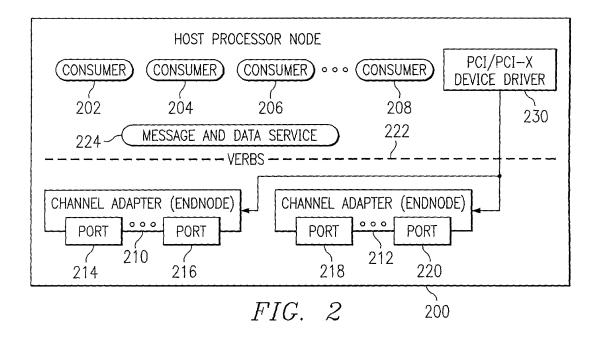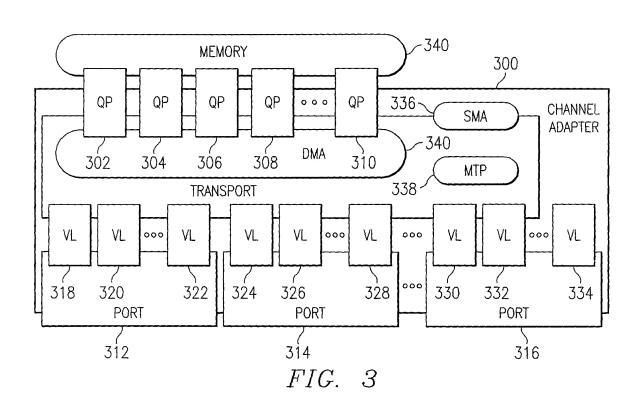AUS920010492US1
Craddock et al.
End Node Partitioning Using Local Identifiers

1/6

FIG. 1

AUS920010492US1
Craddock et al.
End Node Partitioning Using Local Identifiers

2/6

HOST PROCESSOR NODE

(CONSUMER) (CONSUMER) (CONSUMER) ∘ ∘ ∘ (CONSUMER)

202   204   206   208

PCI/PCI-X DEVICE DRIVER

230

224 — (MESSAGE AND DATA SERVICE)   222

— — — VERBS — — —

CHANNEL ADAPTER (ENDNODE)

PORT   ∘ ∘ ∘   PORT

214   210   216

CHANNEL ADAPTER (ENDNODE)

PORT   ∘ ∘ ∘   PORT

218   212   220

FIG. 2   200

MEMORY   340

300

QP   QP   QP   QP   ∘ ∘ ∘   QP   336   SMA   CHANNEL ADAPTER

302   304   306   308   DMA   310   340

338   MTP

TRANSPORT

VL   VL   ∘∘∘   VL   VL   VL   ∘∘∘   VL   ∘∘∘   VL   VL   ∘∘∘   VL

318   320   322   324   326   328   ∘∘∘   330   332   334

PORT   PORT   PORT

312   314   316

FIG. 3

AUS920010492US1
Craddock et al.
End Node Partitioning Using Local Identifiers

3/6

FIG. 4

DATA SEGMENT 1	~444
DATA SEGMENT 2	~446
DATA SEGMENT 3	~448

DATA SEGMENT 4	~438
DATA SEGMENT 5	~440
DATA SEGMENT 6	~442

400

406

CONSUMER

WORK REQUEST
410

WORK REQUEST
412

WORK COMPLETION
414

RECEIVE WORK QUEUE

| WQE | WQE | WQE | WQE |
| 416 | 418 | 419 | 420 |

SEND WORK QUEUE

| WQE | WQE | WQE | WQE |
| 422 | 424 | 426 | 428 |
402

COMPLETION QUEUE

| CQE | CQE | CQE | CQE |
| 430 | 432 | 434 | 436 |
404

408

HARDWARE

FIG. 5

500

MESSAGE DATA

| DATA SEGMENT 1 502 | DATA SEGMENT 2 504 | DATA SEGMENT 3 506 |

508 — PACKET

510

516   518

| ROUTING HEADER | TRANSPORT HEADER | PACKET PAYLOAD | CRC |

514

512
DATA PACKET (ROUTED UNIT OF WORK)

AUS920010492US1
Craddock et al.
End Node Partitioning Using Local Identifiers

4/6

FIG. 6

AUS920010492US1
Craddock et al.
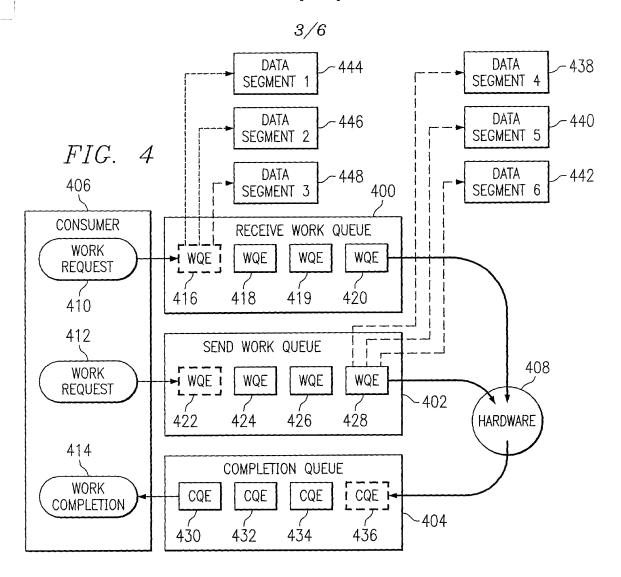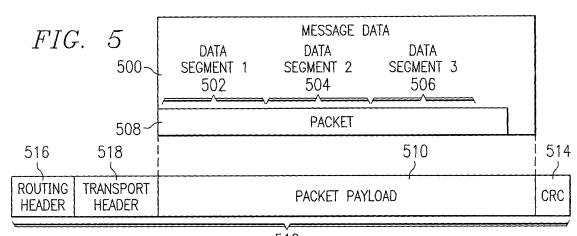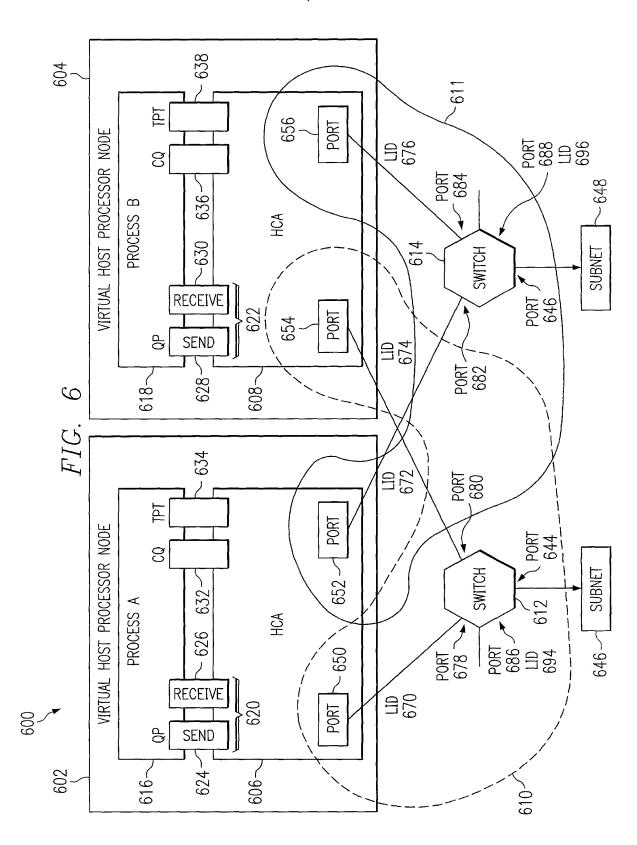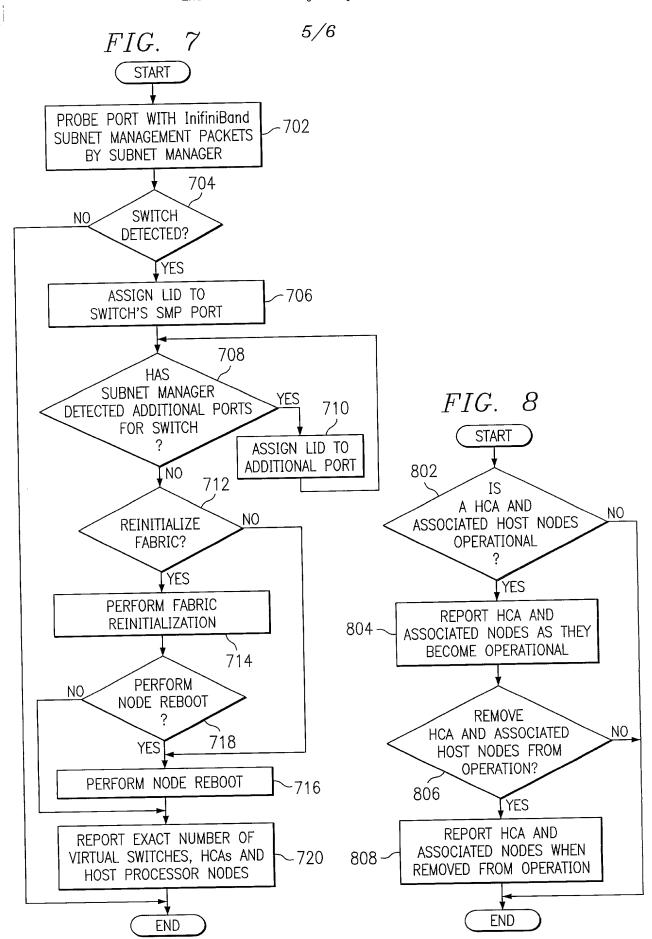End Node Partitioning Using Local Identifiers

5/6

FIG. 7

START

PROBE PORT WITH InifiniBand
SUBNET MANAGEMENT PACKETS ⟋702
BY SUBNET MANAGER

704

SWITCH
DETECTED?   NO

YES

ASSIGN LID TO ⟋706
SWITCH'S SMP PORT

708

HAS
SUBNET MANAGER
DETECTED ADDITIONAL PORTS   YES   710
FOR SWITCH
?

ASSIGN LID TO
ADDITIONAL PORT

NO

712

REINITIALIZE   NO
FABRIC?

YES

PERFORM FABRIC
REINITIALIZATION

714

PERFORM
NODE REBOOT   NO
?

YES   718

PERFORM NODE REBOOT ⟋716

REPORT EXACT NUMBER OF
VIRTUAL SWITCHES, HCAs AND ⟋720
HOST PROCESSOR NODES

END

FIG. 8

START

802

IS
A HCA AND
ASSOCIATED HOST NODES   NO
OPERATIONAL
?

YES

804   REPORT HCA AND
ASSOCIATED NODES AS THEY
BECOME OPERATIONAL

REMOVE
HCA AND ASSOCIATED   NO
HOST NODES FROM
OPERATION?

806   YES

808   REPORT HCA AND
ASSOCIATED NODES WHEN
REMOVED FROM OPERATION

END

AUS920010492US1
Craddock et al.
End Node Partitioning Using Local Identifiers

6/6

START

902 — DOES HCA BECOME OPERATIONAL AT POWER ON? — YES

NO

904 — DOES HCA BECOME OPERATIONAL AFTER POWER ON? — NO

YES

906 — PROBE PORT WITH InfiniBand SUBNET MANAGEMENT PACKETS BY SUBNET MANAGER

908 — SWITCH DETECTED? — YES

NO

910 — ASSIGN LID TO SWITCH'S SMP PORT(S)

912 — HAS SUBNET MANAGER DETECTED ADDITIONAL PORTS FOR SWITCH USING SMP SWEEP? — YES

NO

914 — HAS SUBNET MANAGER DETECTED ADDITIONAL PORTS FOR SWITCH USING A TRAP FROM THE SWITCH INDICATING A STATE CHANGE? — NO

YES

916 — ASSIGN LID TO ADDITIONAL PORT

PERFORM FABRIC REINITIALIZATION? — NO

918

YES

919 — PERFORM FABRIC REINITIALIZATION

920 — PERFORM NODE REBOOT

922 — REPORT HCAs AND HOST PROCESSOR NODES AS THEY BECOME OPERATIONAL

END

FIG. 9